

NES Pilot Design

How the National Employment Service would prove - or disprove - its value for money, and build reusable national infrastructure, before any commitment to national scale. Companion to the [White Paper](#), the [costing](#) and the [evidence annex](#).

Purpose

The pilot exists to do one thing the costing cannot: **convert the Grade-C assumptions that the value-for-money case rests on into Grade-A measured evidence** (the evidence grades run from A, strongest, to D, weakest) - take-up, the wellbeing gain, the value of the work, displacement/additionality, and the rate of transition to unsubsidised work - so that any national-scale decision rests on what the National Employment Service (NES) *actually does*, not on what a spreadsheet assumed. It does so while building reusable national infrastructure that retains value even if the policy does not proceed.

Design principles

- **Voluntary, made-to-pay, default-on;** Real Living Wage; graduated hours. Not workfare, not bolted onto conditionality.
- **Progression-oriented and real-employer-involved by default** for anyone who can move toward the market - the lesson of Germany's ABM job-creation schemes (from the German *Arbeitsbeschaffungsmaßnahmen*): sheltered, dead-end placements *harm* the employable. **Permanent placement reserved** for those who genuinely cannot reach the market (severe disability; collapsed-market regions).
- **No payment-by-results for the placement phase** (the Work Programme cream-skimmed the hardest, who are NES's point); outcome-linked elements, if any, only on sustained post-placement employment.
- **Post-placement transition support** (from the Employment Retention and Advancement (ERA) trial: durable gains need it and take ~5 years to show).
- **Additionality engineered, not asserted:** local commissioning panels + dedicated non-market entities; the funded-post test; union/sector sign-off; no host may cut paid staff and backfill.
- **Evaluation built in before the first participant and pre-registered.**

1. What the pilot must prove (the evidence questions)

Each maps to a Grade-C gap in the costing:

Question	Why it matters	How the pilot answers it
Take-up - what % of the eligible far-from-market pool volunteer?	The 300k national figure is a stipulated ~20%-of-1.5m placeholder	Measured directly against the eligible pool in pilot regions
Wellbeing gain - is it ~0.46 of a WELLBY (a wellbeing-adjusted life-year, the standard unit governments use to value changes in wellbeing), more, or less, for <i>this</i> population?	Largest soft term in the benefit-cost ratio (BCR, the value returned per pound spent)	ONS4 life-satisfaction (the Office for National Statistics' four standard questions), baseline/6/12/24mo, treatment vs control
Value of the work (output)	Anchored only to the Future Jobs Fund's (FJF) employer benefit	Independent valuation of placement output (social-value framework)
Displacement / deadweight	The standing criticism of Kickstart by the National Audit Office (NAO, Parliament's public-spending watchdog); assumed, never measured	Linked employer-employee data + control-area labour-market tracking
Transition rate to unsubsidised work (bridge cohort)	Drives the multi-year societal BCR	Tracked over the 5-year window, no extrapolation
Which cohorts/placements work	The both-by-circumstance design is untested	Segmented outcomes by cohort (bridge / floor-individual / floor-place)

2. The three no-regret assets (the dual-use infrastructure)

Built **lean-but-reusable** - well-architected for reuse, not gold-plated. Each retains national value even if the NES *policy* does not proceed:

- 1. Real-time evaluation automation** - pipelines linking placement, benefit, His Majesty's Revenue and Customs (HMRC) earnings and survey data into a continuously-updated evaluation dashboard. A reusable government-evaluation capability that structurally fixes the FJF failure (evidence arriving after the decision).
- 2. A federated cross-boundary matching layer** - *extending* the Department for Work and Pensions' (DWP) Find a Job (and aligning with its 2026 rebuild - coordinate, don't collide), adding the missing pieces: matching to **paid community/regeneration placements** alongside market vacancies, with the mobility dimension (commute/relocate/remote). Reusable as national employment infrastructure.
- 3. The federated, opt-in, consent-based skills-data layer** - riding existing rails (the Department for Education's (DfE) Learning Records Service, HMRC, DWP) with the citizen consenting to a unified view; broker, not a new national database. Reusable national skills infrastructure in its own right.

3. Scope

- **Regions: 4-5, deliberately mixed** (post-industrial / collapsed-market, coastal, rural, and at least one urban) - to test both the *bridge* logic (where a market exists) and the *permanent-floor* logic (where it doesn't), and to give the staggered-rollout evaluation regional variation. Selection criteria: existing community-org capacity to host; a labour-market shortage NES can feed (e.g. construction for housing); local-authority/mayoral willingness to co-design; ability to establish a clean baseline.
- **Cohorts:** *bridge* (closer to market - progression-oriented, real-employer placements + a Sector-based Work Academy Programme (SWAP)-style fast-track); *floor-individual* (far-from-market / health / disability - **Individual Placement and Support (IPS, a well-tested method of helping people into work) component at "Good"+ fidelity**, ~£3-5k/participant); *floor-place* (collapsed-market regions - dedicated non-market entities, permanent-capable). Needs-assessed at intake, not treated as homogeneous.
- **Scale: ~12,000 paid placements** across the regions at pilot steady state (range 8-15k), sized so the *bridge* cohort has ~5-10k for 80% power to detect a ~5pp (percentage-point) transition effect; plus a larger, cheaper *mission-1 support* throughput (~30-50k/yr). Some regions also run a randomised waitlist where oversubscribed (clean randomised controlled trial (RCT, the gold-standard 'does X cause Y' experiment) variation).
- **Duration: ~3 years of operation** (pilot Y1-2, interim read Y2-3), with the **evaluation tracking continuing to the full 5-year window** - the post-placement tail only appears later, so the funding/evaluation commitment must outlast the placements.

4. The embedded evaluation (pre-registered)

Built in before the first participant starts and pre-registered (the methods, and why each is chosen over the alternatives, are detailed in the Evidence Annex, Part 5):

- **Pre-register** outcomes, comparison-group construction and estimators before launch; protected evaluation budget written into the pilot's mandate (the FJF lesson).
- **Experimental where possible:** randomised waitlist/rationing in oversubscribed regions (clean RCT); **regression discontinuity** at any eligibility-score or duration threshold.
- **Quasi-experimental for the rest: staggered regional rollout** analysed with **Callaway-Sant'Anna (2021)** group-time difference-in-differences (not two-way fixed effects); **synthetic control** for whole-region effects; **entropy balancing**, not propensity-score matching (PSM, a statistical way of comparing scheme participants with similar non-participants) (King & Nielsen 2019), where matching is needed; PSM only as a robustness check.
- **Primary outcomes:** sustained unsubsidised employment (bridge cohort); **wellbeing** (ONS4, a primary outcome not an afterthought); social value of the work; **displacement measured directly** (linked employer-employee data + control areas).
- **Power:** ~5-10k treated per cohort for a 5pp effect; pre-powered before the design is fixed.
- **No extrapolation beyond the 5-year window.**

5. Decision gates (each evidence-gated)

- **Gate 1 - proceed to live pilot:** legal authority and data-sharing gateways (with Data Protection Impact Assessments (DPIAs), the formal checks of privacy risks) in place; the lean infrastructure live; commissioning panels and additionality safeguards operating; the pre-registered evaluation protocol published; baseline measured.
- **Gate 2 - expand within the pilot:** positive interim participant outcomes; no material displacement detected; placement quality and additionality holding; take-up in the expected range.
- **Gate 3 - national scale:** the **BCR proven** on measured (not assumed) wellbeing/output/displacement/transition; stable platform; and a **confirmed cross-departmental funding settlement** (the wrong-pocket problem solved). Absent this, NES does not scale - and the no-regret infrastructure still stands.

6. Cost envelope (value-for-money applies to the pilot too)

Order-of-magnitude, from the per-placement figures in the costing (~£26k gross, ~£15.8k net to the paying department) and the one-off setup:

- **Paid placements:** ~12,000 × ~£26k ≈ **£310m/yr gross** (~£190m/yr net to the paying department after benefit-offset and tax-recapture); over ~2.5 running years, ****~£0.5-0.8bn gross**** cumulative.
- **Mission-1 support:** ~£100-150m/yr (much cheaper per head).
- **One-off no-regret infrastructure (the three assets): ~£150-500m** - retains value regardless of the pilot's verdict.
- **Embedded evaluation:** ~£20-50m over the pilot.
- **Indicative total: ~£1.0-1.8bn over three years** (a "lean pilot" of fewer regions/placements trades statistical power for a lower bill). For scale: comparable to Kickstart (~£1.9bn) and Restart, and **an order of magnitude below the ~£4bn/year national steady state** - while generating the evidence that would justify (or stop) that commitment.

7. Governance and funding

- **Lean national design authority + spine** (platform, standards, evaluation, the make-work-pay reform) with **local delivery** (mayors, councils, dedicated entities) - per the hybrid form.
- **Cross-departmental sponsorship from the outset** (DWP, DfE, Health, Communities) - because the benefits land across departments, the pilot's funding and governance should too; this also pre-builds the Gate-3 funding settlement.
- **Statutory/protected evaluation commitment** so the evidence survives a spending review (the FJF was killed before its evaluation reported).

8. The no-regret payoff (if the pilot disproves the BCR)

If the pilot shows NES does *not* deliver value for money, the country still keeps: a reusable real-time programme-evaluation capability; a federated cross-boundary job-and-community-placement matching

layer (enhancing Find a Job); and a consent-based skills-data layer. The policy bet is de-risked because the infrastructure is useful either way - which is what makes the pilot fundable.

Why these dials - scale and regional mix

- **Scale: the full pilot - ~12,000 paid placements across 4-5 regions, ~£1-1.8bn over three years.**
Chosen over a leaner sub-£1bn version because the result has to be one a Treasury and a sceptical opposition would believe; an underpowered pilot risks a weak-but-positive finding being dismissed.
- **Regional mix: balanced** - collapsed-market + coastal + rural + urban - to test the permanent-floor logic *and* the bridge/transition logic, since the bridge cohort is where the multi-year BCR is strongest and must be proven too.
- The pilot scale (~5-10k in the bridge cohort) is set by the statistical power needed to detect a ~5pp transition effect; the no-regret infrastructure (~£150-500m) is justified independently of the pilot's verdict.